

WEAK-LENSING PEAK FINDING: ESTIMATORS, FILTERS, AND BIASES

FABIAN SCHMIDT¹ AND EDUARDO ROZO^{2,3,4}

¹ Theoretical Astrophysics, California Institute of Technology, M/C 350-17, Pasadena, CA 91125, USA

² Department of Astronomy & Astrophysics, The University of Chicago, Chicago, IL 60637, USA

³ Kavli Institute for Cosmological Physics, Chicago, IL 60637, USA

Received 2010 September 14; accepted 2011 April 11; published 2011 June 24

ABSTRACT

Large catalogs of shear-selected peaks have recently become a reality. In order to properly interpret the abundance and properties of these peaks, it is necessary to take into account the effects of the clustering of source galaxies, among themselves and with the lens. In addition, the preferred selection of magnified galaxies in a flux- and size-limited sample leads to fluctuations in the apparent source density that correlate with the lensing field. In this paper, we investigate these issues for two different choices of shear estimators that are commonly in use today: globally normalized and locally normalized estimators. While in principle equivalent, in practice these estimators respond differently to systematic effects such as magnification and cluster member dilution. Furthermore, we find that the answer to the question of which estimator is statistically superior depends on the specific shape of the filter employed for peak finding; suboptimal choices of the estimator+filter combination can result in a suppression of the number of high peaks by orders of magnitude. Magnification and size bias generally act to increase the signal-to-noise ν of shear peaks; for high peaks the boost can be as large as $\Delta\nu \approx 1-2$. Due to the steepness of the peak abundance function, these boosts can result in a significant increase in the observed abundance of shear peaks. A companion paper investigates these same issues within the context of stacked weak-lensing mass estimates.

Key words: gravitational lensing: weak

Online-only material: color figures

1. INTRODUCTION

The abundance of rare massive objects (clusters) in the universe has emerged as a powerful probe of cosmology (Vikhlinin et al. 2009; Rozo et al. 2010; Mantz et al. 2010; Henry et al. 2009; Vanderlinde et al. 2010). Many different techniques can be used to find these clusters, such as optical identification, X-rays, the Sunyaev–Zel’dovich effect, and gravitational lensing. Among these, lensing stands out as being the least sensitive to the complicated baryonic physics that govern galaxy formation and the non-thermal processes that affect the dynamics of the intra-cluster medium. Consequently, the lensing signal is expected to be the easiest to predict, a fact that has fostered great interest in developing weak-lensing cluster finders (Schneider 1996; Hennawi & Spergel 2005; Hamana et al. 2004) and has even led to the publication of several lensing-selected cluster samples (Miyazaki et al. 2007; Wittman et al. 2006; Gavazzi & Soucail 2007). In practice, lensing selection suffers from significant projection effects: the lensing signal of a cluster can be enhanced by a favorable projection of a triaxial halo, by associated mass distributions (substructure filaments), or by the chance superposition of large-scale structure (LSS) along the line of sight. Fortunately, such superpositions can be calibrated by relying on numerical simulations, and weak-lensing peak statistics remain a promising probe for cosmological physics (Wang et al. 2009; Dietrich & Hartlap 2010; Kratochvil et al. 2010; Marian et al. 2010).

Weak-lensing shear measurements use the shapes of distant galaxies in order to statistically extract the lensing signal. To date, most work has assumed that source galaxies are randomly distributed, whereas in practice we know that galaxies are clustered. The apparent clustering receives two contributions: one

from intrinsic (physical) clustering of the galaxies, henceforth referred to as source clustering; the other from fluctuations in the galaxy density induced by lensing itself via magnification and size bias (Broadhurst et al. 1995; Schmidt et al. 2009b). The intrinsic clustering of source galaxies acts to increase the noise in the shear measurement while the lensing-induced fluctuations can bias the shear measurement if they are not properly accounted for. In fact, the magnification effect on the galaxy density can be seen as a probe of weak gravitational lensing in its own right (Schneider et al. 2000; Van Waerbeke et al. 2010; Rozo & Schmidt 2010).

In a companion paper (Roza et al. 2011, henceforth referred to as Paper I), we study similar issues for stacked weak-lensing analyses of groups and clusters. The two papers are highly complementary as Paper I focuses on high signal-to-noise weak-lensing mass calibration of objects that have been previously identified and then stacked, whereas in this work we focus on identifying shear peaks in the low signal-to-noise regime.

Following Paper I, we discuss two possible filtered shear estimators that differ in the way the estimator is normalized. The first estimator uses a fixed normalization (e.g., Maturi et al. 2005; Wang et al. 2009) and is therefore sensitive to the overall modulation of the source density field by the lensing signal. The second estimator uses an individual normalization for each point in the sky (e.g., Miyazaki et al. 2007). In this approach, the fluctuations in the density of background galaxies are partially canceled out which reduces noise but also shrinks the extra signal due to magnification. Moreover, a location-based normalization estimator can lead to dilution of the lensing signal if the source population is contaminated by cluster galaxies (see Paper I for a more detailed discussion).

Throughout, we adopt a fiducial flat Λ cold dark matter cosmology with $h = 0.7$, $\Omega_m = 0.28$, $n_s = 0.96$, and a power spectrum normalization of $\sigma_8 = 0.85$ at $z = 0$. All masses

⁴ Einstein Fellow.

are defined as M_{200m} , i.e., enclosing an average density of 200 times the mean matter density, and are given in units of $M_\odot h^{-1}$ (the value of h is thus only relevant for the shape of the matter power spectrum). The source galaxies are assumed to follow the redshift distribution expected for the Dark Energy Survey (DES),⁵

$$\frac{dN}{dz} \propto z^2 \exp[-(z/z_0)^\beta], \quad (1)$$

with $z_0 = 0.5$ and $\beta = 1.4$, and we assumed a density of $\bar{n} = 10 \text{ arcmin}^{-2}$.

Section 2 presents the shear estimators and discusses how magnification impacts each of these estimators in turn. We also discuss the choice of filter function and optimal filter scale. Section 3 presents the results on the statistics of lensing peaks. We conclude in Section 4. Details on the lensing calculations and two derivations regarding the variance of smoothed shear filters have been relegated to Appendices A, B, and C.

2. SMOOTHED SHEAR ESTIMATORS

In this section, we focus on estimators of the form

$$\hat{A}(\vec{\theta}) = \frac{1}{\bar{n}} \sum_i e_i W(\vec{\theta}_i - \vec{\theta}), \quad (2)$$

where the sum is over all galaxies,⁶ \bar{n} is the mean source galaxy density, e_i is the tangential component of the ellipticity of galaxy i (with respect to the relative position $\vec{\theta}_i - \vec{\theta}$), and W is an arbitrary filter which we assume is unity normalized,

$$\int d^2\theta W(\vec{\theta}) = 1. \quad (3)$$

For the time being, we leave the filter unspecified. The estimator, Equation (2), has a *fixed* normalization. We consider the alternative choice of a varying normalization in Section 2.3.

In order to derive the statistical properties of \hat{A} , we proceed as follows: first, we divide the sky into infinitesimal pixels of area $\Delta\Omega$ such that the number of galaxies in each pixel is either 0 or 1. Letting $n(\vec{\theta})$ be the galaxy density field on the sky, we can rewrite Equation (2) as

$$\hat{A}(\vec{\theta}) = \frac{1}{\bar{n}} \sum_\alpha n_\alpha e_\alpha W_\alpha \Delta\Omega, \quad (4)$$

where the sum is now over all pixels and the index α denotes that the quantity of interest is evaluated for the pixel centered on $\vec{\theta}_\alpha$. For example, $n_\alpha = n(\vec{\theta}_\alpha)$. We now set

$$n_\alpha = \bar{n}(1 + \delta_\alpha), \quad (5)$$

where δ_α is the galaxy density fluctuation (in this section, we assume that the fluctuations are purely Poisson). Further, we assume $\langle e_\alpha \rangle = g_\alpha$, where $g = \gamma/(1 - \kappa)$ is the reduced tangential shear. In other words, we assume the shear response to be 1 or equivalently consider scaled ellipticities where the

shear response has been taken into account. We then find that the expectation value of \hat{A} is given by

$$\langle \hat{A}(\vec{\theta}) \rangle = \sum_a g_a W_a \Delta\Omega = \int d^2\theta' g(\vec{\theta}') W(|\vec{\theta} - \vec{\theta}'|). \quad (6)$$

For the second equality, we have let $\Delta\Omega \rightarrow 0$ (continuum limit) and correspondingly replaced $\sum_a \Delta\Omega$ with $\int d^2\theta$. Further, we have assumed that the source galaxy overdensity δ is uncorrelated with the shear. In the following discussion, we will set $\vec{\theta} = 0$ without loss of generality. We can compute the variance of \hat{A} in a similar fashion. In particular, using Equations (4) and (5), we find

$$\hat{A}^2 = \sum_{\alpha\beta} (1 + \delta_\alpha \delta_\beta) e_\alpha e_\beta W_\alpha W_\beta (\Delta\Omega)^2, \quad (7)$$

where we have ignored terms proportional to δ since these will become zero upon taking the expectation value. Neglecting any clustering of the source galaxies for the moment, we have

$$\langle \delta_\alpha \delta_\beta \rangle = \frac{1}{\bar{n} \Delta\Omega} \delta_{\alpha\beta}, \quad (8)$$

where $\delta_{\alpha\beta}$ denotes the Kronecker symbol while the expectation value of the galaxy ellipticities takes the form (recall that e_i stands for the tangential component of the ellipticity)

$$\langle e_\alpha e_\beta \rangle = g_\alpha g_\beta + \frac{\sigma_e^2}{2} \delta_{\alpha\beta}. \quad (9)$$

Here, σ_e denotes the rms (intrinsic) galaxy ellipticity of the sample. Using these expressions, taking the expectation value of Equation (7), and subtracting $\langle \hat{A} \rangle^2$, we find

$$\text{Var}(\hat{A}) = \frac{1}{\bar{n}} \int d^2\theta W^2(\vec{\theta}) \left(\frac{\sigma_e^2}{2} + g^2(\vec{\theta}) \right). \quad (10)$$

In many cases, $|g| \ll \sigma_e \sim 0.3$ and the term $g^2(\vec{\theta})$ in Equation (10) is often neglected. In that case,

$$\text{Var}(\hat{A}) = \frac{\sigma_e^2}{2\bar{n}} \int d^2\theta W^2(\vec{\theta}) = \frac{\sigma_e^2}{2\bar{n}} \frac{1}{4\pi\Theta^2}, \quad (11)$$

where for the second equality we have assumed a Gaussian shear filter of width Θ (Equation (12)), which is the standard result for the variance of \hat{A} (van Waerbeke 2000).

2.1. Filter Functions

A variety of smoothing kernels have been proposed for averaging the shear, including top-hat, aperture mass (Schneider et al. 1998), and matched filters (Maturi et al. 2005; Marian & Bernstein 2006). Some differences in the filters are due to the various goals they were designed to achieve; for example, to reduce contributions from small or large scales or the mass-sheet degeneracy present in shear measurements.

In this paper, we will discuss three commonly used or proposed filters, which are shown in Figure 1. It is instructive to compare this figure with Figure 8 in Appendix A, which shows the angular scales contributing to the lensing signal. Our first and perhaps simplest choice is a Gaussian shear filter,

$$W(\theta) = \frac{1}{2\pi\Theta^2} \exp\left(-\frac{\theta^2}{2\Theta^2}\right), \quad (12)$$

⁵ <http://www.darkenergysurvey.org/>

⁶ Note that for simplicity, and in order to keep results general, we have not included any galaxy weights, which in practice will be used in particular if photometric redshifts are available.

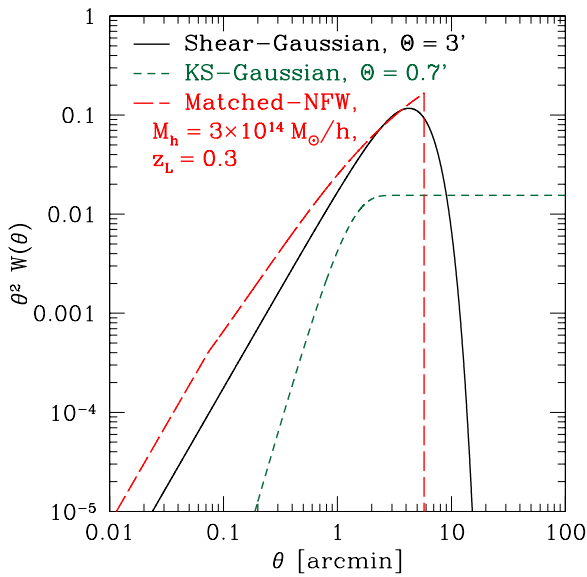


Figure 1. Shear filter functions considered in this paper, as a function of angular separation: Gaussian-smoothed shear filter, Gaussian-smoothed convergence filter using the method of Kaiser & Squires (1993), and matched-NFW filter (Marian & Bernstein 2006).

(A color version of this figure is available in the online journal.)

where Θ denotes the filter scale, which can be chosen to maximize signal-to-noise for a given lens mass and redshift. This filter is also similar in shape to the filter presented in Maturi et al. (2005) designed to reduce the contribution from unassociated LSS.

Another choice is a filter constructed using the method of Kaiser & Squires (1993) to yield an estimator of the Gaussian-smoothed convergence κ (“KS-Gaussian”). This is equivalent to searching for peaks in smoothed convergence maps (e.g., Hamana et al. 2004; Wang et al. 2009, but see caveats below). The shear filter function is given by

$$W(\theta) = \frac{1}{\pi\theta^2} \left[1 - \left(1 + \frac{\theta^2}{2\Theta^2} \right) \exp\left(-\frac{\theta^2}{2\Theta^2}\right) \right], \quad (13)$$

where Θ is the width of the Gaussian convergence filter. This filter is usually used by pixelizing the sky in patches of a few arcmin² area and measuring the average shear in each pixel. The resulting shear map is then convolved with the KS-Gaussian filter. While we do not explicitly consider pixelized estimators here, under certain conditions the globally normalized estimator, Equation (2), is equivalent to a pixelized estimator: this is the case if either the shear in each pixel is estimated with a global normalization (such as in Equation (2)) and pixels are weighted equally or if the shear in a pixel is measured with a local normalization and each pixel is then inverse-variance weighted, i.e., weighted by the number of galaxies in the pixel in the further analysis.

Note that this filter is not capable of being normalized through Equation (3). Instead, the normalization is defined through the convergence filter. Due to the non-local relationship between shear and convergence, this filter is quite different from the Gaussian shear filter. As shown in Figure 1, small scales are downweighted, while in principle arbitrarily large scales are weighted equally ($\theta^2 W \sim \text{const}$). This is a consequence of the mass-sheet degeneracy present in the shear, which can only be broken by including very large scales. We will see that due to these differences, a shear estimator using the KS-Gaussian

filter behaves quite differently than an estimator using the other filters discussed here.

Finally, one can choose a filter matched to the expected signal of a Navarro-Frenk-White (NFW; Navarro et al. 1996) lensing halo. Following Marian & Bernstein (2006), we choose

$$W(\theta) = C g(\theta), \quad (14)$$

where g is the reduced shear profile of an NFW halo (see Appendix A) and the constant is determined from the normalization constraint, Equation (3). Note that as written, the filter is optimal for uniform noise, and in the absence of magnification it can easily be generalized to take into account the effects discussed in this paper. In Marian & Bernstein (2006), the filter was truncated at a scale corresponding to the virial radius (or R_{200}) of the halo (i.e., $W = 0$ for $\theta > \theta_{\text{max}}$); in general, one can vary the truncation scale of the filter. Figure 1 shows that the matched-NFW filter (with truncation at R_{200}) is in fact quite similar to the Gaussian shear filter.

We will return to the choice of filter and optimal filter scale Θ in Section 2.5.

2.2. Magnification and Source Clustering

In the presence of foreground lensing matter, the number density of a flux- and size-limited sample of background galaxies is affected by two effects: galaxies get pushed over the flux and/or size threshold by lensing magnification, and their density is diluted because the observed patch of sky is stretched (Broadhurst et al. 1995). These two effects combine in such a way that the observed source galaxy density field is related to the unlensed background density field via⁷

$$n_{\text{obs}}(\vec{\theta}) = \bar{n}[1 + \delta_g(\vec{\theta})]\mu(\vec{\theta})^{q/2}, \quad (15)$$

where the magnification μ is given by⁸

$$\mu(\vec{\theta}) = \frac{1}{(1 - \kappa)^2 - |\gamma|^2} = 1 + 2\kappa + 3\kappa^2 + |\gamma|^2 + \dots, \quad (16)$$

and q characterizes the contributions from magnification and size bias (Schmidt et al. 2009a, 2009b). Specifically, the parameter q is given in terms of β_f and β_r , the logarithmic slopes of the flux and size distributions, as

$$q = 2\beta_f + \beta_r - 2 \quad (17)$$

$$\beta_f \equiv -\frac{d \ln n_{\text{obs}}}{d \ln f} \Big|_{f=f_{\text{min}}}, \quad \beta_r \equiv -\frac{d \ln n_{\text{obs}}}{d \ln r} \Big|_{r=r_{\text{min}}}. \quad (18)$$

Here f denotes flux and r stands for apparent size of the galaxies. For definiteness, we assume a value of $q = 1.5$ in the middle of the range estimated by Schmidt et al. (2009b). Consider now the expectation value of \hat{A} in the presence of magnification. Inserting Equation (15) into Equation (4) and taking the expectation value, we find

$$\langle \hat{A} \rangle = \int d^2\theta \mu^{q/2}(\vec{\theta}) g(\vec{\theta}) W(\vec{\theta}), \quad (19)$$

⁷ In the weak-lensing literature, it is customary to linearize the magnification term $\mu^{q/2} \approx (1 + q\kappa)$. The parameterization equation (15) is the general expression valid into the moderate lensing regime, see Broadhurst et al. (1995).

⁸ In the following, we will ignore the fact that the source redshift distribution dN/dz itself depends on θ and will always calculate μ assuming the average dN/dz . Since the lensing efficiency varies slowly with redshift, we expect such fluctuations to have negligible impact.

where we have assumed that the lensing field does not correlated with the source density field δ_g . We see that the increase in the number of background sources (assuming $q > 0$) leads to a higher signal, as expected.

We now turn to estimating the variance of \hat{A} in the presence of magnification bias and source clustering. We assume that a weak-lensing shear peak occurs when \hat{A} is evaluated at the center of a lensing halo. For now, we assume that there is no contribution to the lensing shear and magnification from other matter along the line of sight. The clustering of the source population, which we neglected in Section 2, is an additional noise contribution to \hat{A} . Using Equation (15), we have

$$\langle \delta_\alpha \delta_\beta \rangle = \frac{1}{\bar{n} \mu^{q/2} \Delta \Omega} \delta_{\alpha\beta} + \xi_{\alpha\beta}, \quad (20)$$

where $\xi_{\alpha\beta} = \xi(|\vec{\theta}_\alpha - \vec{\theta}_\beta|)$ is the source galaxy angular correlation function. Here, we neglect higher moments of the angular distribution of source galaxies, which can become relevant on very small scales. In order to calculate ξ , we make another approximation: we assume that galaxies follow the matter distribution with a linear bias of ~ 1 and use the fitting formula of Smith et al. (2003) for the nonlinear matter power spectrum together with the redshift distribution, Equation (1) (see also the Appendix of Paper I). Note that ξ is the *observed* correlation function, which in principle is also modified by magnification induced by LSS (Matsubara 2000; Hui et al. 2007). Since this is a small ($< 10\%$) correction to a usually subdominant noise contribution, we neglect this effect here given our rough approximations for ξ . However, we do take into account that the source density and correspondingly the shot noise are modified through magnification. The mean and variance of source ellipticities remain unchanged by magnification, so that as before $\langle e_i \rangle = g_i$ and $\langle e_i e_j \rangle = \frac{\sigma_e^2}{2} \delta_{ij} + g_i g_j$.

Putting everything together and plugging it into Equation (7), we find

$$\text{Var}(\hat{A}) = V_{\text{shot}} + V_{\text{src}} \quad (21)$$

$$V_{\text{shot}} = \frac{1}{\bar{n}} \int d^2\theta \mu^{q/2}(\vec{\theta}) W^2(\vec{\theta}) \left(\frac{\sigma_e^2}{2} + g^2(\vec{\theta}) \right) \quad (22)$$

$$V_{\text{src}} = \int d^2\theta \int d^2\theta' S S' \xi(|\vec{\theta} - \vec{\theta}'|), \quad (23)$$

where

$$S(\vec{\theta}) = \mu^{q/2}(\vec{\theta}) g(\vec{\theta}) W(\vec{\theta}), \quad (24)$$

and primed and un-primed variables are evaluated at $\vec{\theta}'$ and $\vec{\theta}$, respectively. We have split the variance of \hat{A} into a shot-noise contribution V_{shot} , and a contribution from source clustering V_{src} . Comparing Equation (21) with Equation (10), we see that magnification increases the shot noise in \hat{A} due to the increased source density (but not as fast as it increases the value of \hat{A} itself). In addition, the clustering of source galaxies adds to the variance of \hat{A} .

Figure 2 shows the two contributions to the variance of \hat{A} for a Gaussian filter (Equation (12) in Section 2.1) as a function of the filter scale Θ . The thick lines show results including magnification ($q = 1.5$), while the thin lines are without magnification. For the results with magnification, we have assumed that the estimator is centered on an NFW lensing

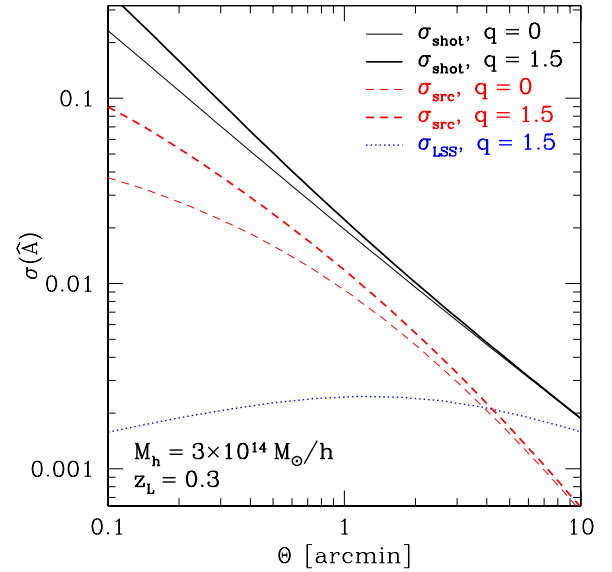


Figure 2. Contributions to the variance of smoothed shear estimators as a function of filter scale Θ for a Gaussian shear filter: shot noise $\sigma_{\text{shot}} = \sqrt{V_{\text{shot}}}$ (Equation (22)), source clustering noise $\sigma_{\text{src}} = \sqrt{V_{\text{src}}}$ (Equation (23)), and large-scale structure variance σ_{LSS} (Appendix B). σ_{shot} and σ_{src} are shown with (thick lines; $q = 1.5$) and without (thin lines; $q = 0$) magnification, while σ_{LSS} is only shown including the very small magnification correction (again for $q = 1.5$).

(A color version of this figure is available in the online journal.)

halo of mass $3 \times 10^{14} M_\odot h^{-1}$ at redshift $z_L = 0.3$. We see that source clustering is subdominant compared to shot noise for an average source density of $\bar{n} = 20 \text{ arcmin}^{-2}$. For higher source densities, source clustering will become important. We also see that lensing bias increases both sources of noise, with V_{src} being affected more strongly.

Finally, we consider one more source of variance for shear estimators: that from the lensing field induced by LSS itself. Generally, massive dark matter halos or density peaks reside in overdense regions. This associated LSS can add to or subtract from the lensing signal of the halo itself. Unfortunately, calculating this contribution properly is only possible using N -body simulations (especially when taking into account magnification). What we can calculate, however, is the estimated variance of \hat{A} induced by uncorrelated LSS, σ_{LSS} , as detailed in Appendix B (see also Hoekstra 2001). The result is shown as the dotted line in Figure 2, again for the Gaussian shear filter. Clearly, LSS noise is subdominant for this filter as long as Θ is not very large; still, for $\Theta \gtrsim 3 \text{ arcmin}$, it cannot be neglected. Furthermore, the magnitude of σ_{LSS} depends on the type of filter chosen: for the KS-Gaussian filter, the LSS noise is much more significant, $\sigma_{\text{LSS}} \sim 0.01$ for $\Theta \lesssim 1 \text{ arcmin}$.

2.3. Another Shear Estimator

An alternative to Equation (2) as the choice of smoothed shear estimator uses a location-based normalization:

$$\hat{B}(\vec{\theta}) = \frac{\sum_i e_i W(\vec{\theta}_i - \vec{\theta})}{\sum_i W(\vec{\theta}_i - \vec{\theta})}, \quad (25)$$

where both sums run over all galaxies. The normalizing denominator removes some of the fluctuations in the source galaxy density (which can be intrinsic or survey specific, such as varying the depth of the observations). The statistical properties of

this type of estimator have been studied in detail by Lombardi & Schneider (2001, 2002). We can write Equation (25) as

$$\hat{B}(\vec{\theta}) = \frac{\hat{A}(\vec{\theta})}{\hat{N}(\vec{\theta})}, \quad (26)$$

$$\begin{aligned} \hat{N}(\vec{\theta}) &= \frac{1}{\bar{n}} \sum_{\alpha} n_{\alpha} W_{\alpha} \Delta\Omega \\ &= \int d^2\vec{\theta}' \mu^{q/2}(\vec{\theta}') [1 + \delta(\vec{\theta}') W(\vec{\theta} - \vec{\theta}'), \end{aligned} \quad (27)$$

where in the last line we have employed the continuum limit.

It is worth pointing out some differences between \hat{A} and \hat{B} and their response to systematic effects due to source clustering. For instance, fluctuations in the number density of galaxies *behind* the lens are canceled out in \hat{B} , while they contribute to \hat{A} . On the other hand, if there is an overdensity of galaxies associated with the lens or in the foreground (present in the sample, for example, due to uncertainties in photometric redshifts), these galaxies systematically reduce the value of \hat{B} since they contribute zero shear to the numerator in \hat{B} , despite being included in the denominator. This so-called dilution (Bernardeau 1998; Medezinski et al. 2007) does not directly affect the estimator \hat{A} . A detailed discussion of the different systematics induced in both estimators by photometric redshift uncertainties can be found in Paper I.

When calculating the statistics of \hat{B} , we face the problem that $\langle \hat{A}/\hat{N} \rangle \neq \langle \hat{A} \rangle / \langle \hat{N} \rangle$ (see also Paper I). However, we can employ this approximation if the effective number of galaxies used in the shear estimator, $\sim \bar{n}\Theta^2$ where Θ is the filter scale, is much larger than unity. In this case, for the expectation value we have

$$\langle \hat{B} \rangle = \frac{\langle \hat{A} \rangle}{\langle \hat{N} \rangle} = \frac{\int d^2\theta g(\vec{\theta}) \mu^{q/2}(\vec{\theta}) W(\vec{\theta})}{\int d^2\theta \mu^{q/2}(\vec{\theta}) W(\vec{\theta})}, \quad (28)$$

where we have again assumed that the shear is uncorrelated with source density. We see that the effect of magnification on \hat{B} is partially canceled by the denominator, and therefore the expectation value for \hat{B} is only weakly dependent on magnification corrections. In the absence of magnification and systematics, $\langle \hat{B} \rangle = \langle \hat{A} \rangle$.

When calculating the variance of \hat{B} , however, it is necessary to take into account the covariance between numerator and denominator. The full expression for $\text{Var}(\hat{B})$ is derived in Appendix C and is given in Equation (C4). The gist of it is that fluctuations in the number of source galaxies, due to both shot noise and source clustering, are partially canceled out (but see below). This becomes especially important for high masses where the source clustering contribution to the variance begins to dominate. Finally, the LSS variance is the same for both estimators, assuming that it is dominated by the weak-lensing regime (lowest order in κ, γ).

The top panel of Figure 3 shows the average expected signal-to-noise ν , defined as $\nu = \langle \hat{A} \rangle / [\text{Var}(\hat{A})]^{1/2}$ for both \hat{A} and \hat{B} , as a function of halo mass. For comparison, the bottom panel of Figure 3 shows the cumulative number of halos above mass M (per square degree) in a narrow redshift slice centered at the lens redshift $z_L = 0.3$. We have again assumed an NFW lensing halo and show results both for a Gaussian shear filter of width $\Theta = 3'$ and a KS-Gaussian filter with $\Theta = 0.7'$ (the choices

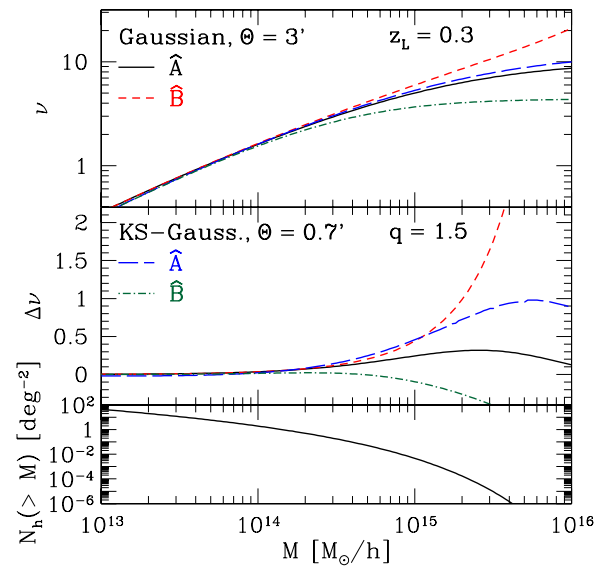


Figure 3. Top panel: signal-to-noise of estimators \hat{A} and \hat{B} without magnification for a lens halo of mass M at $z_L = 0.3$, using a Gaussian shear filter and a KS-Gaussian filter (see Section 2.1). Middle panel: change in signal-to-noise of \hat{A} and \hat{B} induced by magnification, for the same filters and lensing halo. Bottom panel: cumulative number of halos above mass M per square degree in a redshift slice centered at $z = 0.3$ and with width $\Delta z = 0.1$.

(A color version of this figure is available in the online journal.)

of Θ will be justified in Section 2.5). Further, we have ignored magnification for the moment. For the Gaussian filter, we find that \hat{B} is superior to \hat{A} in terms of signal-to-noise for massive halos. For lower mass halos, where $g^2 \ll \sigma_e^2$, both estimators are equivalent. Similar conclusions hold for the NFW-optimized filter.⁹

The KS-Gaussian filter shows a very different behavior: here, \hat{B} has significantly less signal-to-noise than \hat{A} at all relevant masses, while \hat{A} performs very similarly to the corresponding Gaussian estimator. The reason is that the covariance between the numerator and denominator of \hat{B} , which leads to the partial cancelation of fluctuations for the other shear filters, is strongly suppressed for a KS-Gaussian filter. This is a consequence of the inclusion of very large scales in the filter (see Appendix C). Without significant covariance between numerator and denominator, \hat{B} is just the ratio of two noisy quantities and not surprisingly has less signal-to-noise than \hat{A} . As mentioned in Section 2.1, however, this filter is commonly used on a pixelized shear map rather than directly on the galaxies. Hence, Equation (25) is actually not used with the KS-Gaussian filter in practice. Nevertheless, this result illustrates an important point: the choice of estimator (e.g., \hat{A} versus \hat{B}) and filter function W has to be done jointly, as the two are interrelated.

2.4. Impact of Magnification

We now turn to the impact of magnification on \hat{A} and \hat{B} . Magnification changes both the value (signal) of smoothed shear estimators as well as the signal-to-noise. The latter effect is shown in the middle panel of Figure 3. Let us discuss the Gaussian filter first. We see that for $M \lesssim 3 \times 10^{15} M_{\odot} h^{-1}$,

⁹ In Paper I, we found that both types of estimators are statistically equivalent. This is because, first, source clustering is negligible for the thin annuli considered in the stacked weak-lensing context and second, for the relevant radial scales, the stacked analysis is in the regime of $g^2 \ll \sigma_e^2$.

magnification tends to increase the signal-to-noise in both estimators, an effect that increases with mass. The effect is in fact greater for \hat{B} than for \hat{A} , even though the effect on the value of \hat{B} is smaller. This is because magnification acts to decrease the variance of \hat{B} in two ways: first, the increased source density reduces shot noise, and second, magnification increases the covariance between numerator and denominator, improving the cancelation of fluctuations in the galaxy number above that without magnification. In case of \hat{A} , we see a turnover of the magnification effect at very high masses. This is because for very high halo masses (which are necessary to produce this high signal-to-noise), source clustering takes over as the dominant source at this filter scale. While the source clustering noise scales similarly with the lensing quantities as the signal itself ($\propto \mu^{q/2}g$), it is weighted by W^2 rather than W ; this weighting favors smaller radii where the lensing magnification is stronger (see Figure 8 in Appendix A), thus boosting the noise more than the signal. Note, however, that for such halos the optimal filter scale is significantly larger than the assumed 3 arcmin. On the other hand, fluctuations due to source clustering are largely canceled out in \hat{B} , and we do not see this reversal for this estimator. Instead, the boost in signal-to-noise grows strongly toward larger halo masses.

Again, the KS–Gaussian filter behaves differently: a significant increase in $\nu(\hat{A})$ is seen, while the effect on $\nu(\hat{B})$ is a small decrease. The causes for the differences are, for \hat{A} , that source clustering is somewhat smaller for the KS–Gaussian filter than for the Gaussian filter, and for \hat{B} , that magnification increases the variance of the estimator slightly faster than it increases the value of \hat{B} itself.

It is also interesting to consider what impact magnification has on the mass that one would assign to each shear peak. Assuming one can predict a relation $\hat{A} = \hat{A}(M)$, then from the amplitude of the shear peak one can estimate its mass. If, however, one fails to account for the boost in signal due to magnification, one will systematically overestimate the corresponding halo mass.¹⁰ The systematic mass offset is approximately given by

$$\frac{\Delta M_{\text{vir}}}{M_{\text{vir}}} = \frac{d \ln M}{d \ln \hat{A}} \frac{\Delta \hat{A}}{\hat{A}}, \quad (29)$$

where the logarithmic slope $d \ln \hat{A} / d \ln M \approx 0.5\text{--}0.7$ depending on mass, and correspondingly for \hat{B} . The top panel of Figure 4 (thick lines) shows the relative bias $\Delta M/M$ obtained with a Gaussian filter with $\Theta = 3'$. Clearly, for the most massive halo biases as large as tens of percent are possible. Note that the bias in \hat{B} is *smaller* than that of \hat{A} . This is because we are only relying on the amplitude of \hat{B} to estimate a cluster's mass, and as we have seen, the amplitude of \hat{B} is less affected than that of \hat{A} . The thin lines in the top panel of Figure 4 show the expected bias if one would instead use the signal-to-noise of a lensing peak to estimate the halo mass, calculated using a similar relation to Equation (29), but for $\nu(\hat{A})$, $\nu(\hat{B})$ instead \hat{A} , \hat{B} . Here, the situation is reversed at high masses: the mass bias is larger for \hat{B} than for \hat{A} , reaching order unity at $M \approx 10^{15} M_{\odot} h^{-1}$. Moreover, the bias in \hat{A} turns over at very high masses (for a

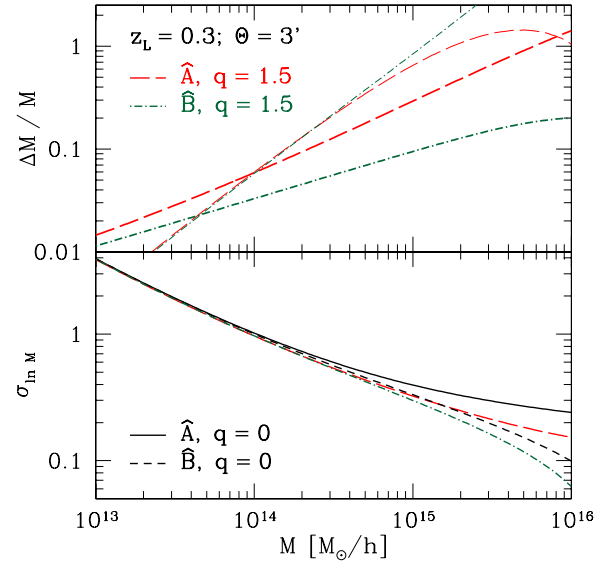


Figure 4. Estimated mass bias when neglecting magnification (top) and relative error (bottom) on the mass measurement using filtered shear estimators \hat{A} and \hat{B} . In the top panel, the thick lines show the mass bias expected if the signal itself is used to estimate the halo mass (thick lines), and if the signal-to-noise is used for the estimate (thin lines). In all cases, we have assumed a lens redshift of $z_L = 0.3$ and a Gaussian shear filter with $\Theta = 3'$.

(A color version of this figure is available in the online journal.)

fixed filter scale). This reflects the behavior of the magnification effect on ν seen in Figure 3.

Turning now to the statistical uncertainty in the recovered mass, we have seen that magnification helps increase the signal-to-noise of a given halo. Thus, we expect properly accounting for lensing bias will reduce the statistical error in cluster mass estimates. The expected error in log-mass can be estimated as

$$\sigma_{\ln M} = \frac{d \ln M}{d \ln \hat{A}} \frac{1}{\nu_A}. \quad (30)$$

This is shown in the bottom panel of Figure 4. Again, the improvement in mass resolution becomes relevant at the high-mass end. We also see how the better statistical performance of \hat{B} over \hat{A} for this filter reflects in the smaller mass uncertainty at the high-mass end. Note that this should only be seen as a rough estimate; in practice, halo triaxiality and projection effects will significantly increase the error in the recovered masses.

2.5. Choice of Filter Scale

In order to optimize the filter shape, one has to adopt a metric with which different filters can be compared. In the following, we will focus on the goal of maximizing the signal-to-noise for a given lensing halo, which is most directly relevant to shear peak counts.

In order to test the relative performance of the different filters, we show the signal-to-noise ν for a $3 \times 10^{14} M_{\odot} h^{-1}$ lensing halo at fixed lens redshifts of $z_L = 0.3, 0.6$ in Figure 5, as a function of the filter scale Θ . In the case of the matched-NFW filter, $\Theta = \theta_{\text{max}}$ is the truncation scale of the filter. We use \hat{A} in all cases. The optimal signal-to-noise is determined by a balance between the signal \hat{A} , which decreases with increasing filter size, and the noise which also decreases. As expected, the optimal filter scale depends on the redshift of the lensing halo and shifts to smaller scales for higher z_L . It also depends on the lens mass.

¹⁰ Note that in order to predict a mass–shear-peak relation one needs to assume a lens redshift as well as a profile shape, i.e., halo concentration. Here, we assume that the true values of z_L and c are known.

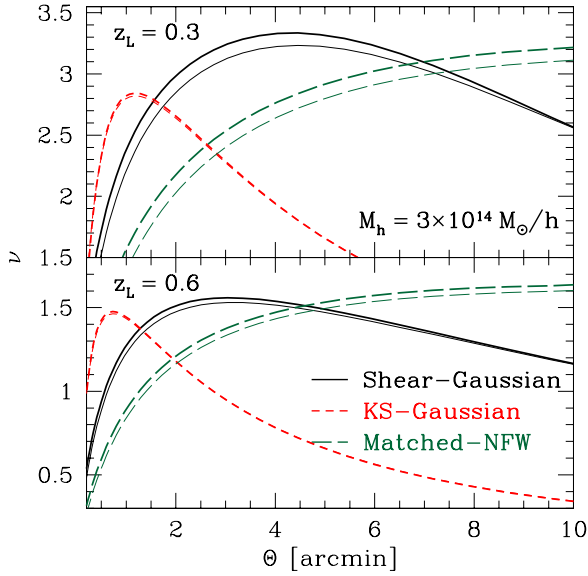


Figure 5. Signal-to-noise of the estimator \hat{A} as a function of filter scale for different filter shapes (truncation scale in case of the matched-NFW filter), for a $3 \times 10^{14} M_{\odot} h^{-1}$ halo at two different redshifts: $z_L = 0.3$ (top panel) and $z_L = 0.6$ (bottom panel). The thin lines show results without magnification, while the thick lines include magnification ($q = 1.5$).

(A color version of this figure is available in the online journal.)

We see that all three filters peak at roughly the same signal-to-noise for \hat{A} , though the Gaussian shear filter seems to slightly outperform the other filter shapes for the $z_L = 0.3$ case. Furthermore, the filter scale at which this peak is reached is very different: it is approximately a factor of four smaller for the KS-Gaussian as compared to the Gaussian filter. In all cases, the reduction in signal-to-noise beyond the peak value is due to the growing importance of LSS noise. This reflects the fact that the likelihood of a chance superposition with unrelated matter concentrations increases with the filter scale (Hoekstra 2003; Maturi et al. 2005).

The relative importance of the three noise contributions, shot noise, source clustering, and LSS noise, depends on the filter shape. In the case of the KS-Gaussian filter, the LSS noise is much more significant, a factor of ~ 10 higher at the optimal filter scale than for the other two filters. Again, this is a consequence of the inclusion of very large scales in this filter. This will also be of relevance to the contribution of correlated LSS to smoothed shear estimators.

Figure 5 also shows the effect of magnification (thick lines versus thin lines): for this halo, magnification boosts the peak signal-to-noise by $\sim 5\%$ for the Gaussian and NFW-optimized filters. In case of the KS-Gaussian, the effect is slightly smaller, $\sim 3\%$, due to the preference for large scales in that filter. Note that the optimal filter scale is moved to slightly smaller values by magnification. This is because the signal in \hat{A} increases more steeply with decreasing filter size when including magnification.

Before deciding on an optimal filter, however, it is necessary to additionally take into account the effect of *associated* LSS along the line of sight to the lens. Furthermore, a purely Gaussian calculation of the uncorrelated LSS noise as used here is likely not sufficient. Both of these effects will tend to make a true optimal filter narrower than suggested by Figure 5. We will return to this and a discussion of further potential systematic issues in Section 4.

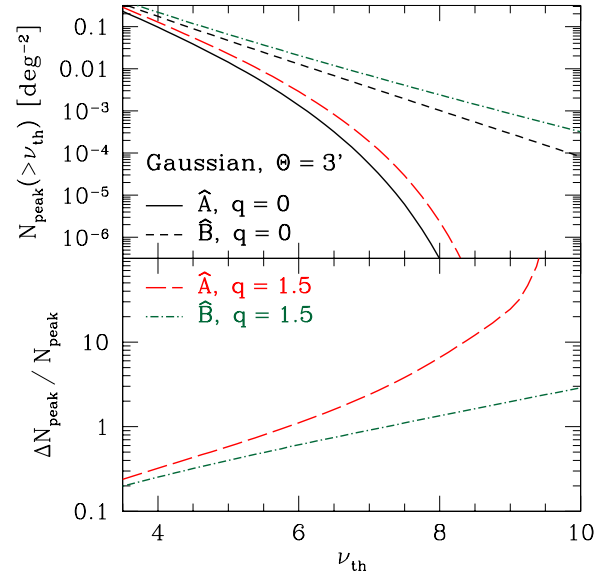


Figure 6. Upper panel: average number peak of peaks $N_{\text{peak}}(> \nu_{\text{th}})$ (per deg^2) above the signal-to-noise threshold ν_{th} , with and without lensing bias for the two estimators \hat{A} , \hat{B} , using a Gaussian shear filter with $\Theta = 3$ arcmin. Lower panel: relative magnification effect on the peak counts, $\Delta N_{\text{peak}}(> \nu_{\text{th}})/N_{\text{peak}}(\nu_{\text{th}})$.

(A color version of this figure is available in the online journal.)

3. THE PEAK FUNCTION

We now turn to investigating the impact of magnification and filter choice on the abundance of detected shear peaks. To do so, we assume that each shear peak corresponds to an NFW halo and that the peak position is the halo center. While in practice one expects some fraction of weak-lensing peaks to arise due to chance superpositions of multiple halos, the results we obtain concerning how magnification impacts weak-lensing peak finding should be indicative of the whole population.

We estimate the average number of lensing peaks within a solid angle Ω_s above a given signal-to-noise threshold ν_{th} as follows:

$$N_{\text{peak}}(> \nu_{\text{th}}) = \Omega_s \int_0^\infty \frac{c dz_L}{H(z_L)} \chi^2(z_L) \times \int_{\ln M_{\min}(z_L, \nu_{\text{th}})}^\infty d \ln M \frac{dn}{d \ln M}(z_L, M), \quad (31)$$

where c is the speed of light, $H(z)$ is the Hubble expansion rate, $\chi(z)$ is the comoving distance to redshift z , $dn/d \ln M$ is the halo mass function, and M_{\min} is defined via

$$\nu(M_{\min}, z_L) = \nu_{\text{th}}. \quad (32)$$

Hence, magnification enters by lowering the effective z -dependent mass threshold of the survey, so that the number of peaks is increased. We use the fitting function of Tinker et al. (2008) to calculate $dn/d \ln M$ as a function of mass and redshift from the linear matter power spectrum (recall that $M = M_{200\text{m}}$ throughout).

Figure 6 shows the peak count statistics $N_{\text{peak}}(> \nu_{\text{th}})$ (peaks per deg^2) with and without magnification using a Gaussian filter with fixed filter scale $\Theta = 3$ arcmin. We also show the relative magnification effect $\Delta N_{\text{peak}}/N_{\text{peak}}$. Evidently, estimator \hat{B} is more efficient at finding peaks for this filter, with or without lensing bias: the number of high signal-to-noise peaks

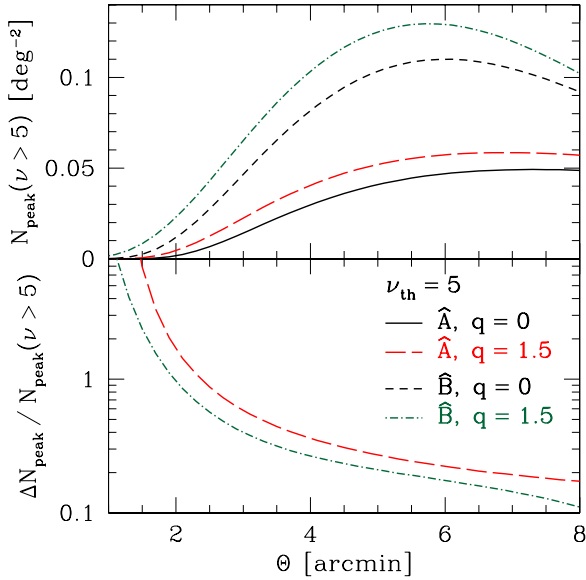


Figure 7. Upper panel: average number peak of peaks $N_{\text{peak}}(\nu > 5)$ (per deg^2), with and without magnification for the two estimators \hat{A} , \hat{B} as a function of the filter scale Θ . Other survey specifications as in Figure 2. Lower panel: relative magnification effect on the peak counts, $\Delta N_{\text{peak}}(\nu > 5)/N_{\text{peak}}(\nu > 5)$. (A color version of this figure is available in the online journal.)

($\nu \gtrsim 6$) is higher than that found by \hat{A} by orders of magnitude. Further, magnification can boost the weak-lensing peak counts significantly for both \hat{A} and \hat{B} , by more than a factor of two for high peaks.

Note that for very high peaks, i.e., massive halos, a filter scale of 3 arcmin is not optimal. Hence, it is also interesting to consider the number of peaks above a fixed threshold as a function of the filter width (Figure 7). The different number of peaks here reflects the different mass thresholds for different filter scales at a fixed signal-to-noise threshold. Again, the location-normalized estimator \hat{B} yields more peaks than \hat{A} , and magnification increases the number of 5σ shear peaks found with either estimator by $\sim 30\%$. Note that since lensing bias pushes the optimal filter scale to smaller values, the number of peaks increases faster with scale when incorporating the impact of magnification. Not surprisingly, we find that choosing smaller filters can increase the relative importance of magnification significantly (see the lower panel of Figure 7).

Finally, we note that the result of the comparison \hat{A} versus \hat{B} reverses for the KS–Gaussian filter: \hat{A} performs far better for this filter, as expected after Section 2.3.

4. SUMMARY AND DISCUSSION

We have compared the expected signal-to-noise of the two most common types of filtered shear estimators for different filter functions and studied how magnification impacts these estimators. In our signal-to-noise considerations, we also include the effects of the intrinsic clustering of source galaxies, and the variance of the estimators due to uncorrelated LSS. The former noise contribution becomes important for massive lensing halos, while the latter’s importance only depends on the filter shape and scale considered.

We find that estimator and filter function need to be chosen jointly and cannot be regarded as independent. For example, for the Gaussian shear filter, the location-normalized estimator

(\hat{B}) is statistically superior to the globally normalized estimator (\hat{A}) for high peaks. This is because fluctuations in the number of galaxies are canceled out to first order in \hat{B} . For the KS–Gaussian filter on the other hand, the situation is reversed: \hat{A} performs far better than \hat{B} . According to our (certainly not exhaustive) results, a location-normalized estimator with a Gaussian-type shear filter appears to perform best statistically. The question of optimal filter+estimator combination deserves more attention, as the abundance of high peaks can be suppressed by orders of magnitude for suboptimal choices (Figure 6).

Another finding, of equal importance to their statistical properties, is that the two types of estimators respond very differently to uncertainties in the photometric redshifts (see the discussion in Paper I). Specifically, the globally normalized estimator \hat{A} does not suffer from the so-called dilution effect affecting \hat{B} and should be much less sensitive to contamination of the source sample by galaxies associated with the lens.

Turning to magnification, we find that it affects both of the estimators we considered, and for all filter functions. While the signal-to-noise of either estimators can be boosted significantly (up to $\Delta\nu \sim 1$ –2), the value of the estimator \hat{A} is generally affected more strongly than that of \hat{B} . Indeed, if one is to estimate halo masses based solely on their smoothed shear signal, halo mass overestimates as large as tens of percent are possible. Not surprisingly, the increase in signal-to-noise especially of high peaks also results in comparable boosts to the abundance of observed peaks. The magnitude of the effect depends strongly on the filter scale as well: smaller filter scales lead to a much larger boost in signal due to magnification, pushing the optimal filter scale toward smaller values.

Among the possibly relevant systematic issues we have not considered here are the impact of scatter in the halo profiles (concentration), halo triaxiality, and correlated LSS. Similar to how an optical survey preferentially picks out compact rather than extended faint galaxies, weak-lensing peaks are biased toward more concentrated halos as well as triaxial halos whose long axis is oriented along the line of sight. Correlated LSS can also boost an otherwise undetected lensing halo above the detection threshold. While magnification generally increases the lensing signal, the increase itself is proportional to the lensing signal and will hence increase this selection toward concentrated or favorably projected halos. A detailed study of this effect is beyond the scope of this paper.

In weak-lensing peak finding, the signal is generally dominated by scales larger than the scale radius of halos. We thus expect that the effect of baryons on the halo density profiles will be a subdominant systematic.

In light of these results, it is clear that a proper modeling of magnification as well as source clustering will be a necessary component of cosmological interpretations of the shear peak function. Furthermore, these effects should also be taken into account when designing an optimal estimator+filter combination for shear peak finding. Fortunately, incorporating these two effects is fairly straightforward, both for analytic calculations and numerical studies with N -body simulations. Magnification, by increasing the signal-to-noise of the shear signal, has the potential to significantly boost the statistical power of shear-selected samples of objects. Thus, properly accounting for magnification and size bias should allow us to maximize the cosmological potential of shear peak statistics.

We thank Richard Ellis, Richard Massey, James Taylor, and Scott Dodelson for helpful discussions.

F.S. is supported by the Gordon and Betty Moore Foundation at Caltech. E.R. is funded by NASA through the Einstein Fellowship program, grant PF9-00068.

APPENDIX A

LENSING BY AN NFW HALO

In this appendix, we review our lens model used for the numerical results. The NFW density profile is given by (Navarro et al. 1997)

$$\rho(r) = \frac{\rho_s}{r/r_s(1+r/r_s)^2}, \quad \rho_s = \frac{M}{4\pi r_s^3 [\ln(1+c) - c/(1+c)]}, \quad (\text{A1})$$

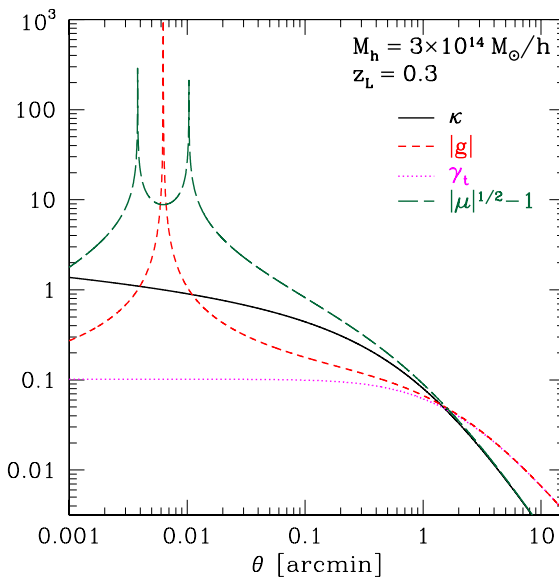
where r_s is the scale radius and $c = R_{200}/r_s$ is taken from the fit of Bullock et al. (2001). Note that this halo profile is not truncated at R_{200} . Actual halo profiles lie somewhere in between the extrapolated NFW profile adopted here and a truncated profile. The differences however appear mainly around R_{200} , which is typically larger than the filter size we will consider. Hence, the details of the outer halo profile do not change the results significantly.

Further, we make the small angle approximation. The lensing quantities κ and γ for an NFW halo have been derived in the literature (Oaxaca Wright & Brainerd 1999; Maturi et al. 2005). For reference, the convergence at the scale radius is given by

$$\kappa_s = \frac{\rho_s r_s}{\Sigma_{\text{cr}}} = \frac{3}{2} \frac{\rho_s r_s H_0^2}{\rho_{\text{cr},0} c^2} (1+z_L) W_L(z_L). \quad (\text{A2})$$

Here, z_L is the halo redshift, $\rho_{\text{cr},0}$ is the critical density today, c is the speed of light, and the lensing weight function is given by

$$W_L(z_L) = \frac{c}{H(z_s)} \int_{z_L}^{\infty} dz_s \frac{\chi_L}{\chi(z_s)} [\chi(z_s) - \chi_L] \frac{dN}{dz_s}, \quad (\text{A3})$$



where $\chi_L = \chi(z_L)$, and dN/dz is the normalized redshift distribution of the source galaxies, Equation (1), determined from a fit to the expected redshift distribution of galaxies for DES (J. Annis 2009, private communication).

Figure 8 (left panel) shows the profile of κ , γ , g as well as μ for a typical detectable lensing halo at $z_L = 0.3$. While κ , γ are well behaved, both g and μ show the strong-lensing caustics, as expected. For this halo and source/lens redshifts, the strong-lensing regime defined by $g(\theta) \sim 1$ is relevant for $\theta \lesssim 0.05$. This regime cannot be treated by weak-lensing analysis techniques, so it should be removed. Fortunately, the bulk of the signal-to-noise for weak lensing is at much larger radii. The right panel of Figure 8 shows the relevant lensing quantities weighted by the Gaussian filter, Equation (12), for $\Theta = 3$ arcmin. Clearly, the bulk of the signal comes from $\theta \sim 0.3$ to 3 arcmin, well outside the strong-lensing regime for this halo. Although not of practical relevance, very massive halos do appear in the calculations where the strong-lensing regime extends to beyond 0.2 arcmin. For those cases, we cap the convergence at $\kappa = 0.5$. This is merely done to lead to convergence of the calculation and only occurs for such high halo masses that it does not impact the results presented here.

APPENDIX B

LARGE-SCALE STRUCTURE VARIANCE

Here we re-derive the variance of \hat{A} due to LSS in the survey (see Hoekstra 2001). Since the bulk of the survey area will have small convergence, we can work in the weak-lensing limit, in which case \hat{A} and \hat{B} are equivalent. Furthermore, we only calculate the leading contribution from the power spectrum of the lensing field, neglecting higher order moments which in principle can become relevant on very small scales. The computation is most conveniently done in Fourier space. For this, we first rewrite Equation (6) in the weak-lensing limit ($g = \gamma$) as an expression for the filtered convergence κ (e.g.,

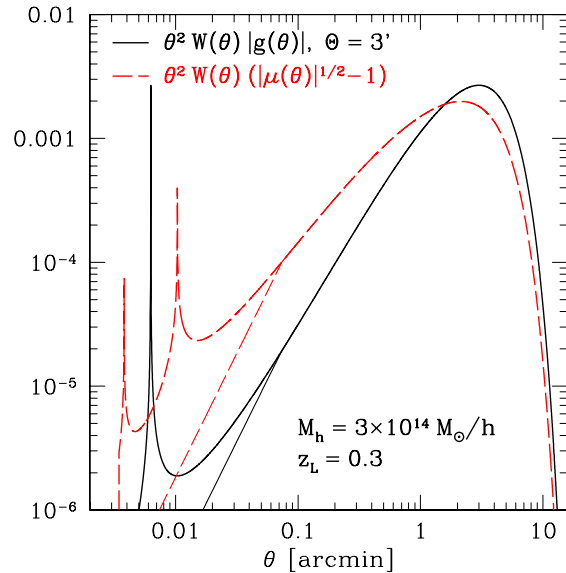


Figure 8. Left panel: profile of convergence κ , tangential shear γ , tangential reduced shear g , and scaled magnification $\sqrt{\mu} - 1$ around a halo of mass $3 \times 10^{14} M_{\odot} h^{-1}$ located at $z_L = 0.3$. Right panel: reduced shear and magnification weighted by area and the Gaussian filter W , for the same parameters as in the left panel. This shows which scales contribute to the signal in Equations (2) and (25). The thin lines show the results capped at $\kappa = 0.5$ (see the text).

(A color version of this figure is available in the online journal.)

Bartelmann & Schneider 2001):

$$\hat{A}(\vec{\theta}) = \int d^2\theta' \kappa(\vec{\theta}') W^\kappa(|\vec{\theta} - \vec{\theta}'|) \quad (\text{B1})$$

$$W^\kappa(\theta) = 2 \int_0^\infty \frac{d\theta'}{\theta'} W(\theta') - W(\theta). \quad (\text{B2})$$

Note the non-local relationship between the shear filter and the corresponding convergence filter. Since κ is a scalar quantity, Equation (B1) can be straightforwardly written in Fourier space:

$$\tilde{A}(\vec{\ell}) = \tilde{\kappa}(\vec{\ell}) \tilde{W}^\kappa(\ell), \quad (\text{B3})$$

where tilded quantities stand for Fourier transforms:

$$\tilde{X}(\vec{\ell}) = \int d^2\theta e^{i\vec{\ell}\cdot\vec{\theta}} X(\vec{\theta}). \quad (\text{B4})$$

Then, using the definition of the angular power spectrum of the convergence

$$\langle \kappa(\vec{\ell}) \kappa(-\vec{\ell}') \rangle = (2\pi)^2 \delta_D(\vec{\ell} - \vec{\ell}') C^\kappa(\ell), \quad (\text{B5})$$

we can write the first-order variance of \hat{A} as

$$\sigma_{\text{LSS}}^2 = \langle \hat{A}^2 \rangle = \int \frac{d^2\ell}{(2\pi)^2} |\tilde{W}^\kappa(\ell)|^2 C^\kappa(\ell). \quad (\text{B6})$$

This quantity is shown as the dotted line in Figure 2, for a Gaussian shear filter, Equation (12) (note that W^κ for this filter is not a Gaussian). It is also straightforward to calculate the leading magnification correction to σ_{LSS} , following the second-order correction to $C^\kappa(\vec{\ell})$ presented in Schmidt et al. (2009a). While we include this correction, the relative change of σ_{LSS} only amounts to a few percent for filter scales $\Theta \gtrsim 1'$ (see also White 2005).

APPENDIX C

VARIANCE OF THE LOCATION-NORMALIZED ESTIMATOR

In this section, we derive the variance of \hat{B} due to shot noise and source clustering, taking into account the covariance between numerator and denominator. We expand $\hat{B} = \hat{A}/\hat{N}$ around its expectation value, $\langle \hat{B} \rangle = \langle \hat{A} \rangle / \langle \hat{N} \rangle$, assuming that fluctuations are much smaller than unity (justified if $\bar{n} \Theta^2 \gg 1$). In close analogy to the derivation in the Appendix of Paper I, this yields

$$\frac{\text{Var}(\hat{B})}{\langle \hat{B} \rangle^2} = \frac{\text{Var}(\hat{A})}{\langle \hat{A} \rangle^2} + 3 \frac{\text{Var}(\hat{N})}{\langle \hat{N} \rangle^2} - 4 \frac{\text{Cov}(\hat{A}, \hat{N})}{\langle \hat{A} \rangle \langle \hat{N} \rangle}, \quad (\text{C1})$$

where

$$\begin{aligned} \text{Cov}(\hat{A}, \hat{N}) &\equiv \langle \hat{A} \hat{N} \rangle - \langle \hat{A} \rangle \langle \hat{N} \rangle = \frac{1}{\bar{n}} \int d^2\theta W^2(\theta) \mu^{q/2}(\theta) g(\theta) \\ &+ \int d^2\theta \int d^2\theta' W \mu^{q/2} W' \mu'^{q/2} g' \xi(|\vec{\theta} - \vec{\theta}'|) \end{aligned} \quad (\text{C2})$$

$$\begin{aligned} \text{Var}(\hat{N}) &= \frac{1}{\bar{n}} \int d^2\theta W^2(\theta) \mu^{q/2}(\theta) \\ &+ \int d^2\theta \int d^2\theta' W \mu^{q/2} W' \mu'^{q/2} \xi(|\vec{\theta} - \vec{\theta}'|). \end{aligned} \quad (\text{C3})$$

Here, we have used the shorthand notation introduced after Equation (24). In each of these equations, the first term denotes the shot-noise contribution, while the second is the contribution from source clustering. We see from Equation (C1) that there is a cancelation of noise terms (both shot noise and source clustering) due to the positive covariance between \hat{A} and \hat{N} . The cancelation is not perfect, since the various terms involve different integrals over shear and magnification. However, for relatively low-mass halos for which $g^2 \ll \sigma_e^2$, the shot-noise term of $\text{Var}(\hat{A})$ (Equation (22)) dominates in Equation (C1). In this limit, we recover $\text{Var}(\hat{B}) = \text{Var}(\hat{A})$ (see Figure 3).

If we neglect magnification and source clustering in both \hat{B} and \hat{A} , so that $\langle \hat{A} \rangle = \langle \hat{B} \rangle$, Equation (C1) simplifies to

$$\text{Var}(\hat{B}) = \text{Var}(\hat{A}) + \frac{\langle \hat{A} \rangle^2}{\bar{n} \mathcal{A}} \left(3 - 4 \frac{\mathcal{A} \int d^2\theta W^2 g}{\int d^2\theta W g} \right), \quad (\text{C4})$$

where we have defined $\mathcal{A} = 1/\int d^2\theta W^2$ as the effective filter area. If the shear g was constant, then the last term in Equation (C4) would evaluate to 4, canceling the g^2 term in $\text{Var}(\hat{A})$ (Equation (22)) and leading to $\text{Var}(\hat{B}) = \sigma_e^2/(2\bar{n} \mathcal{A})$. Equation (C4) makes it clear that the last term determines whether \hat{B} performs better or worse than \hat{A} : if $\mathcal{A} \int W^2 g / \int W g$ is order unity, $\text{Var}(\hat{B}) < \text{Var}(\hat{A})$. If it is much less than 1, $\text{Var}(\hat{B})$ can be significantly larger than $\text{Var}(\hat{A})$. The latter is the case for the KS–Gaussian filter, where $\mathcal{A} \int W^2 g / \int W g \approx 0.2\text{--}0.3$, due to the preferential weighting of very large scales. Similar reasoning applies to the source clustering contribution. This explains why \hat{B} performs worse than \hat{A} for the KS–Gaussian filter. Note that this effect will be mitigated when employing this filter on a pixelized shear map: since the sum in Equation (25) now only runs over galaxies within a small pixel (with $W = 1$) instead of running over all galaxies, there will be significant covariance between numerator and denominator, reducing the noise in the pixelized shear.

REFERENCES

- Bartelmann, M., & Schneider, P. 2001, *Phys. Rep.*, **340**, 291
 Bernardeau, F. 1998, *A&A*, **338**, 375
 Broadhurst, T. J., Taylor, A. N., & Peacock, J. A. 1995, *ApJ*, **438**, 49
 Bullock, J. S., Kolatt, T. S., Sigad, Y., Somerville, R. S., Kravtsov, A. V., Klypin, A. A., Primack, J. R., & Dekel, A. 2001, *MNRAS*, **321**, 559
 Dietrich, J. P., & Hartlap, J. 2010, *MNRAS*, **402**, 1049
 Gavazzi, R., & Soucail, G. 2007, *A&A*, **462**, 459
 Hamana, T., Takada, M., & Yoshida, N. 2004, *MNRAS*, **350**, 893
 Hennawi, J. F., & Spergel, D. N. 2005, *ApJ*, **624**, 59
 Henry, J. P., Evrard, A. E., Hoekstra, H., Babul, A., & Mahdavi, A. 2009, *ApJ*, **691**, 1307
 Hoekstra, H. 2001, *A&A*, **370**, 743
 Hoekstra, H. 2003, *MNRAS*, **339**, 1155
 Hui, L., Gaztañaga, E., & LoVerde, M. 2007, *Phys. Rev. D*, **76**, 103502
 Kaiser, N., & Squires, G. 1993, *ApJ*, **404**, 441
 Kratochvil, J. M., Haiman, Z., & May, M. 2010, *Phys. Rev. D*, **81**, 043519
 Lombardi, M., & Schneider, P. 2001, *A&A*, **373**, 359
 Lombardi, M., & Schneider, P. 2002, *A&A*, **392**, 1153
 Mantz, A., Allen, S. W., Rapetti, D., & Ebeling, H. 2010, *MNRAS*, **406**, 1759
 Marian, L., & Bernstein, G. M. 2006, *Phys. Rev. D*, **73**, 123525
 Marian, L., Smith, R. E., & Bernstein, G. M. 2010, *ApJ*, **709**, 286
 Matsubara, T. 2000, *ApJ*, **537**, L77
 Maturi, M., Meneghetti, M., Bartelmann, M., Dolag, K., & Moscardini, L. 2005, *A&A*, **442**, 851
 Medezinski, E., et al. 2007, *ApJ*, **663**, 717
 Miyazaki, S., Hamana, T., Ellis, R. S., Kashikawa, N., Massey, R. J., Taylor, J., & Refregier, A. 2007, *ApJ*, **669**, 714
 Navarro, J. F., Frenk, C. S., & White, S. D. M. 1996, *ApJ*, **462**, 563

- Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, *ApJ*, **490**, 493
- Oaxaca Wright, C., & Brainerd, T. G. 1999, arXiv:astro-ph/9908213
- Rozo, E., & Schmidt, F. 2010, arXiv:1009.5735
- Rozo, E., Wu, H., & Schmidt, F. 2011, *ApJ*, **735**, 118 (Paper I)
- Rozo, E., et al. 2010, *ApJ*, **708**, 645
- Schmidt, F., Rozo, E., Dodelson, S., Hui, L., & Sheldon, E. 2009a, *ApJ*, **702**, 593
- Schmidt, F., Rozo, E., Dodelson, S., Hui, L., & Sheldon, E. 2009b, *Phys. Rev. Lett.*, **103**, 051301
- Schneider, P. 1996, *MNRAS*, **283**, 837
- Schneider, P., King, L., & Erben, T. 2000, *A&A*, **353**, 41
- Schneider, P., van Waerbeke, L., Jain, B., & Kruse, G. 1998, *MNRAS*, **296**, 873
- Smith, R. E., et al. 2003, *MNRAS*, **341**, 1311
- Tinker, J., Kravtsov, A. V., Klypin, A., Abazajian, K., Warren, M., Yepes, G., Gottlöber, S., & Holz, D. E. 2008, *ApJ*, **688**, 709
- Vanderlinde, K., et al. 2010, *ApJ*, **722**, 1180
- van Waerbeke, L. 2000, *MNRAS*, **313**, 524
- Van Waerbeke, L., Hildebrandt, H., Ford, J., & Milkeraitis, M. 2010, *ApJ*, **723**, L13
- Vikhlinin, A., et al. 2009, *ApJ*, **692**, 1060
- Wang, S., Haiman, Z., & May, M. 2009, *ApJ*, **691**, 547
- White, M. 2005, *Astropart. Phys.*, **23**, 349
- Wittman, D., Dell’Antonio, I. P., Hughes, J. P., Margoniner, V. E., Tyson, J. A., Cohen, J. G., & Norman, D. 2006, *ApJ*, **643**, 128